Application of Propensity Score Theory to the Measurement of the Effects of Military Service

Benjamin A. Fairbank

Operational Technologies Corp.

for

**Contracting Officer's Representative
Hyder Lakhani**

DTIC
SELECTED
MAY 2 5 1995
G

**Organization and Personnel Resources Research Unit
Paul A. Gade, Chief**

**Personnel and Training Systems Research Division
Zita M. Simutis, Director**

April 1995

19950522 090

**United States Army
Research Institute for the Behavioral and Social Sciences**

Approved for public release; distribution is unlimited

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction
of the Deputy Chief of Staff for Personnel

Edgar M. Johnson
Director

Accesion For		
NTIS CRA&I	☒	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

NOTICES

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1995, April	Final Jun 93-Dec 94

4. TITLE AND SUBTITLE

Application of Propensity Score Theory to the Measurement of the Effects of Military Service

5. FUNDING NUMBERS

SBIR Contract
MDA903-93-C-0111

6. AUTHOR(S)

Benjamin A. Fairbank

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Operational Technologies Corp.
4100 N.W. Loop 410, Suite 230
San Antonio, TX 78229-4253

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Institute for the
Behavioral and Social Sciences
ATTN: PERI-RP
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

ARI Research Note
95-35

11. SUPPLEMENTARY NOTES

COR: Hyder Lakhani
This report is submitted to DTIC per requirement as a final contract report. It has not been edited and will not be published as an ARI report.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release;
distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This report focuses on the suitability or propensity score theory to determine the effects of military service on the later life of a Service participant. The limitations inherent in the non-experimental determination of the effects of service have previously precluded strong determination of cause, in part because of the bias introduced by self-selection into military service. Those who serve differ from those who do not serve in at least two ways: they have served, and they have chosen to serve. To attribute any differences later in life to the first of those variables while ignoring the second is not defensible. The present selection presents a method of simulating the phenomena so modeled, then illustrates the simulation with a sample execution. The output of the simulation is examined to determine whether plausible values of the effect of service in the output variables might reasonably be expected to be detected. The differences built into the simulation were recovered, but were not statistically significant.

14. SUBJECT TERMS

Propensity Score   Determination of Cause
Effects of Military Service

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

Summary

This report focuses on the testing of the suitability of propensity score theory to determine the effects of Military Service on the later life of a Service participant. The limitations inherent in the non-experimental determination of the effects of Service have previously precluded strong determination of cause, in part because of the bias introduced by self-selection into Military Service. Those who serve differ from those who do not serve in at least two ways: they have served, and they have chosen to serve. To attribute any differences later in life to the first of those variables while ignoring the second is not defensible. The present paper describes the application of a propensity score model of the effects of self-selection, presents a method of simulating the phenomena so modeled, then illustrates the simulation with a sample execution. The output of the simulation is examined to determine whether plausible values of the effect of Service in the output variables might reasonably be expected to be detected. The differences built into the simulation were recovered, but were not statistically significant.

THIS PAGE INTENTIONALLY LEFT BLANK

The Application of Propensity Score Theory to the Measurement of the Effects of Military Service

I. INTRODUCTION

Objective

The aim of the present investigation has been to determine the feasibility of using propensity score theory as a means of studying the long-term effects of Military Service on the life of individuals who served. While studies of the effects of Military Service have been carried out many times in the past, the results have always suffered to some extent from the limitations inherent in a non-experimental determination of causation. As will be asserted below, new methodological developments have led to the possibility of overcoming some of the difficulties inherent in conducting and interpreting such studies.

Observational studies in the behavioral and social sciences have traditionally been more fruitful as a means of gathering data upon which causal hypotheses may be based than they have been as a methodology for use in testing such hypotheses. The testing of causal hypotheses is generally the strength of the experimental method. There are, however, many important areas of inquiry which are outside of the valid application of experimentation. The social and behavioral sciences are particularly full of areas in which experimentation is impossible for reasons pertaining to ethics, practicality, or the demands of social policy. It is not possible, however desirable it may be, to perform an experiment to determine the effects of Army Service.

Past Findings

Laurence, Ramsberger, and Gribben (1989)[1] performed an extensive investigation of the extent to which Military Service affects the later life of those who are chosen from the population of marginally qualified or unqualified candidates. Two experiments, one of which was intended from its outset to have experimental aspects (Project 100,000) and one unintended (arising as a consequence of a misnorming of the Armed Services Vocational Aptitude Battery, the test used to determine the intellectual qualifications of applicants for Service) were examined in detail to determine the extent to which Service affects a wide spectrum of variables measured at a later date. Their conclusions, based on a study of a large number of variables, was that Service does not have beneficial

---

[1]Laurence, J.H., Ramsberger, P.F., & Gribben, M.A. (1989). Effects of military experience on the post-Service lives of low-aptitude recruits: Project 100,000 and the ASVAB misnorming. (Technical Report 89-29) Alexandria, VA: Human Resources Research Organization.

3

effects. Others, however, have questioned the validity of their conclusions owing to weaknesses in the control of variables on which those who served and who did not serve differed at the time of enlistment (Sticht, 1992)[2]. Further questions have been raised in additional analyses of data closely related to that analyzed by Laurence et al. Sticht, Armstrong, Hickey, and Caylor (1987)[3] have documented them in a study which reaches conclusions other than those of Laurence et al. These studies addressed the effects of Military Service on those with low or marginal intellectual aptitude for military training, but the greater question of interest is whether enlistees who are qualified for Service benefit from such Service.

A number of studies have addressed that question and the general pattern of results has shown that Service does result in benefits to those who have served. Martindale and Poston (1979)[4] showed that Service veterans have higher earning patterns than nonveterans, although evidence from recent Service (Vietnam era) appears to have been associated with smaller and more variable gains than Service longer in the past (WWII and Korea). Similar findings were adduced by Villamez and Kasarda (1976)[5]. A recent major study of Daymont and Andrisani (1986)[6] has shown generally positive economic effects of Service, especially when a longer rather than a shorter post-Service time perspective is utilized. They also report that the economic gains for minority Service members are greater than the gains reported for the nonminority members. Many other studies have been generally in concert with these, although some, such as Crane and Wise (1987)[7], found that Vietnam era veterans actually had costs, rather than benefits, in terms of later earnings. Other investigators have looked into other dependent variables.

---

[2]Sticht, T.G. (1992). How Military Service helped low-aptitude, economically disadvantaged young men of the mid-1960's escape poverty. Research note Number 1. San Diego: Applied Behavioral and Cognitive Sciences, Inc.

[3]Sticht, T.G., Armstrong, W.B., Hickey, D.T., & Caylor, J.S. (1987). Cast-off youth: Policy and training methods from the military experience. New York: Prager.

[4]Martindale, M. & Poston, D.L. (1979). Variations in veteran/nonveteran earnings patterns among World War II, Korea, and Vietnam War cohorts. Armed Forces and Society, 5, 291-243.

[5]Villamez, W.J. & Kasarda, J.D. (1976). Veteran status and socioeconomic attainment. Armed Forces and Society, 2, 407-420.

[6]Daymont, T.N. & Andrisani, P.J. (1986). The economic returns to Military Service. (Technical report USARECSR 86-11) Fort Sheridan, Illinois: U.S. Army Recruiting Command.

[7]Crane, J.R. & Wise, D.A. (1987). Military Service and the civilian earnings of youths. In D.A. Wise (Ed.) Public Sector Payrolls. Chicago: University of Chicago Press. 119-137. (as cited in Lakhani, 1994).

Brumagim and Daymont (1989)[8] examined the perceived attractiveness of the educational benefits derived from Service. They report that the benefits have differential effects for minorities and for nonminorities. Minority enlistees report in greater numbers than do nonminority enlistees that the educational benefits were an important factor in their enlistment decision, but minorities who had reported that educational benefits were important as an incentive were less likely than nonminorities to report satisfaction in attaining the anticipated benefits.

Lakhani (1994 (preliminary version))[9] investigated the correlates of Service in the National Guard and Reserve. He investigated a broad spectrum of variables and found that Military Service (among reservists) contributed to higher civilian income, and that family income was also higher for prior active duty reservists than for non-active duty reservists.

Difficulties of Traditional Methodology

While most of the studies cited above, and many others, show generally positive findings, they do not carry the force that an experiment would in establishing the causative effect of Military Service. They establish unequivocal differences in later life when one compares those who served to those who did not serve, but they do not show that the <u>cause</u> of the difference was Military Service. There remains the possibility that those who serve in the military are in some way different from those who do not serve, not as a consequence of Service, but because of differences that existed before Service. As long as Military Service is subject to self-selection, as has been the case since the inception of the All Volunteer Force (AVF), the possibility of that difference may not be ignored in a consideration of methodologies. We can not, in other words, rule out the possibility that those who chose to serve would have been different twenty years later <u>even if they had not served</u>. As the next section shows, experimentation is the usual way to address such concerns, and in cases where it is impractical or impossible to conduct experiments, there is an alternative approach. That approach is propensity score methodology.

---

[8]Brumagim, A.L. & Daymont, T.N. (1989). The role of military educational benefit programs: Impacts on minority opportunities. <u>Industrial Relations Research Association 42nd Annual Proceedings. 315-325</u>.

[9]Lakhani, H. (1994). <u>The socioeconomic effects of Military Service: Reserve/Guard</u>. Paper presented at the 69th annual Western Economic Association International Conference, Vancouver, Canada, 1994. (Preliminary version).

Propensity Score Theory

The establishment or demonstration of causation within the empirical sciences generally requires that three conditions be met. If we are to assert that Cause C brings about Effect E, then we must be able to demonstrate (1) that C precedes E, (2) that C and E are correlated or that they covary in ways other than correlation (not perfectly, but significantly), and (3) that all other explanations for the occurrence of E may be ruled out. It is the last requirement that is so hard to meet in nonexperimental sciences. Philosophers of science generally state that only the application of the experimental method, with random assignment to groups or with repeated measures, can unambiguously establish causation. Purists in that regard hold that inherently unmanipulable variables, such as sex, race, or eye color can not be considered as causes in the sense of scientific causality.

If, then, one were to try to establish that Military Service were the cause of a later life difference in the circumstances between those who had served and those who had not served, the best way to do so would be by establishing an experimental group and a control group, one of which then served and the other of which did not serve. Assignment to the two groups would have to be made at random. While the draft lottery of the late Vietnam War era came close to meeting those conditions, it was not a perfect experiment, and, in any case, with the advent of the AVF, it quickly became outdated. It is appropriate to note also that "Military Service" is a very general term, and that to state that it might be the cause of a later difference between veterans and nonveterans in, for example, income, is to gloss over a variety of other problems which are primarily associated with construct validity. The establishment of a causal relationship by means of an experiment establishes only so-called internal validity -- the cause may appropriately be identified. The mechanisms which lead to the effect are not specified.

If Military Service were shown to be causally related to increased income (or other variable), it would still not be clear what aspect of Military Service had been responsible. It might be that military training was responsible, or that the social skills learned in the military had been responsible for the effect, or perhaps that the self-discipline that arises as a consequence of learning to accept the discipline of others brought about the effects. Such questions, in any case, go beyond our purpose here. The are not, however, negligible.

Propensity score methodology may be used to compensate in part for the lack of rigorous experimental control in certain classes of investigations (Rosenbaum & Rubin, 1984)[10]. While the details of the application of propensity score methodology are complex and vary from situation to situation,

---

[10]Rosenbaum, P.R. & Rubin, D.R. (1984). Reducing bias in observational studies using subclassification on the propensity score. Journal of the American Statistical Association, 79, 516-524.

the overall idea behind it is not complex. This investigation addresses the circumstance in which self-selection makes direct comparison of those who served and did not serve impossible. Those who did serve and who did not serve do not resemble samples arrived at by random assignment to groups. For purposes of discussion, consider the simplified situation in which one is faced with determining whether an outcome variable, chosen to be income, is or is not affected by a dichotomous variable, in this case, Army Service.

When a predictor variable (or, more realistically, a number of intercorrelated variables) affects both the outcome variable and the tendency to be in one of two groups (in this case, electing to serve in the Military), it is possible to use matching on the predictor variable to compensate, at least partially, for the lack of experimental control (i.e., random assignment). Remember that random assignment to groups would assure that (at least for large samples) the two groups, Service and not Service, will be alike in all respects other than those associated with Service. If it can be shown that predictor variables can "explain" or account for the self-selection decision, then matching the samples of those who served and who did not serve can help to overcome the effects of self-selection. The matching procedure would use the predictor variable to determine the equivalence of the two groups. In particular, if logistic regression is used to model the decision to enlist, then for each person in an enlistment sample and each person in a non-enlistment sample a value may be calculated which expresses the logistic of his or her likelihood to enlist. The modeling of the enlistment decision by the logistic, rather than by a simple probability, is advantageous primarily for methodological reasons. For example, probabilities are strictly bounded by 0 and 1, while the logistic has a valid range from minus infinity to plus infinity (although in practical applications values are almost always between -5 and +5). If a member of a group has a value of a logistic function associated with him or her, then the predicted probability of enlistment (p) can be calculated from the relationship

$$p = 1 / (1 + \exp(-\text{logistic})).$$

It is then possible to assign to each member of each sample a propensity score which reflects enlistment likelihood. Following such assignment, members of the two samples can be matched on their likelihood to enlist (i.e. their propensity), in order to control in part for differences which may account for the self-selection into Service or into the civilian labor force.

The variables chosen to compute the regression coefficients for the propensity score should not be those on which it is possible to carry out exact matching. Sex and race, for example, can be matched exactly and so should be the variables on which to base separate analyses. Continuous variables which predict enlistment propensity, however, are suitable for inclusion in the analyses.

An Experimental Analogy

It may be helpful to consider propensity score theory in an analogy with a true experiment. By considering an "ideal" experiment and seeing why it is impossible to conduct such an experiment, it may be possible to determine the advantages to be gained from the use of propensity score methodology. If it were possible to conduct an experiment on the effects of Military Service, one might select 1000 individuals who wished to join the Service and 1000 who did not wish to join. One would assign, at random, half of the members of each group to enlist in the Service, and forbid the others to serve. One would then wait 20 years for the data to mature, and then go back and determine the dependent variable values for each of the 2000 participants. A situation such as the following would then be the result:

	Those who served	Those who did not serve
Those who wanted to serve	Cell 1	Cell 2
Those who did not want to serve.	Cell 3	Cell 4

By looking at the marginals (i.e. the row and column totals), one can separately determine the effects of Service and the effects of wanting to serve, with the measure of each separated from the influence of the other. Moreover, one can also determine if there is an interaction between wanting to serve and serving. The impossibility of performing such an experiment, however, limits one to observing only cells one and four. Cells two and three will always remain empty, except for such small and theoretically uninteresting cases as those who seek to serve but who are determined to be unqualified. As a consequence, it is possible neither to separate out the effects of Service from those of wanting to serve nor to look for any interaction.

Consider now the idea of a propensity score. If one can analyze the choice of serving or not serving, and model it, then one can approximately fill cells two and three by selecting those who were unlikely to serve but who did so anyway, and those who were likely to serve but did not do so. Usually, more accurate prediction is better than less accurate prediction, but in this case note that the goal is not excellent prediction, just good prediction. If prediction is too good, then using the predictors will have the same disadvantages as using the fact of enlistment. If it is possible to find variables that predict enlistment, then by using those variables to control for whether one enlisted or not, it should possible to disentangle the effects of serving and choosing to serve. Note, however, that

the disentanglement will neither be perfect nor absolutely assured. If a reasonable degree of prediction of enlistment decision is obtained, perhaps an accounting of 25 % of variance, that still leaves 75% of the variance unaccounted for. It may be factors in that 75% of the variance which account for the dependent variable. There is, moreover, no way to know if the unaccounted variance in enlistment decisions is unexplainable error variance or is responsible for the observed effects of Service. It seems necessary to assume that whatever variables are chosen to predict Service would be representative of those that account for the dependent variable. That may be justifiable if it is possible to show a relationship between the predictors and the dependent variable, but even in that case it is not possible to demonstrate that all relevant variables have been found. If one included three more variables among the predictors, is it not possible that the correlation between the predictors and the dependent variable might be higher still? This is especially problematic when we realize that the propensity score might correlate even more highly with the outcome than with enlistment.

Consider now a modification of the experiment above. The modified experiment, now might now take the following form:

	Those who served	Those who did not serve
Those with high propensity scores	Cell 1	Cell 2
Those with low propensity scores.	Cell 3	Cell 4

Now it is possible to fill all of the cells, but at an interesting cost. The cost is that random assignment to cells is no longer possible, so one can not consider this to be a true experiment. The propensity variable can now more properly be seen as a covariate and can be so treated in an analysis. The Service variable can be entered into the analysis as a predictor in an equation. We thus have the equation

Service effect = error + constant + a*propensity + b*Service.

What is the nature of the error? It is not possible to say. It may still be correlated with the desire to serve, but it may not be also. However, to the extent to which propensity correlates with the dependent variable, this represents an advance. Furthermore if propensity does not correlate with the dependent variable, but does correlate with the enlistment decision, what advance is represented? Only the knowledge that either the investigator has missed the relevant variables governing the enlistment decision, or that there is a true effect of Service.

THIS PAGE INTENTIONALLY LEFT BLANK

II. METHODS

Research Plan

In order to demonstrate the feasibility of using the propensity score model for the measurement of the effects of Military Service, a number of simulations of the effects which the propensity score model was designed to detect were to be undertaken. One method of justifying the application of a new methodological approach to the detection of subtle differences (such as difference in later life between those who serve in the Army and those who do not) is to simulate a data set and then apply the new methodology in order to determine whether the new approach is able to detect the differences. If the new method is able to detect the differences, that does not guarantee that similar differences in operationally obtained data (i.e. nonsimulated data obtained from samples of populations of operational interest to the Military) would be detected, but if the differences are not detected, then it is an indication that the method may not be suitable for use on operational data. Thus the detection of constructed differences in simulated data sets is a necessary but not sufficient condition for justifying the use of a new technique.

If the results of the simulations indicate that the method is able to detect the differences built into the simulated data set, then the simulations may be expanded to a Monte-Carlo investigation to determine the power of the proposed tests. A Monte-Carlo investigation of the power of a method is designed to determine the probability of detecting differences of various magnitudes. The power of a statistical test is defined as the probability that the test will detect a difference if there really is a difference. It is equivalent to one minus beta, where beta is the probability of making a type two error, accepting the null hypothesis when it is false. Power can be determined analytically for many statistical tests, especially ones based on the t- and F-distributions, but power analyses have apparently not been undertaken for propensity score analyses. The advantage of using Monte-Carlo investigative techniques as an initial test of a new technique is that the true magnitude of the effect is known, and the sensitivity of the detection method can be assessed. When a new technique is applied initially to data which have been obtained from samples of populations which are of interest to the investigators, then results are difficult to interpret, especially if the results are negative.

The Monte-Carlo tests can be used to estimate or predict the power of the tests in their application to operationally collected data. The initial exploration via simulations, then, was to indicate the feasibility of detailed simulations (i.e. Monte-Carlo investigations) of the effects of Service, and the simulations were to indicate the feasibility of undertaking analyses of operational data.

Simulations

The simulation of the effect of Military Service described here was designed to incorporate the aspects of the effects of Military Service which one might study by means of propensity score methodology. The simulation presented here is simplified to emphasize the aspect of the method which would be most effective in disentangling the effects of Service from the effects of variables which lead one to go into Service (the self-selection effect). The simulation generates a number of values for each simulee in turn. Each pass through the simulation routine adds another simulated member, called a simulee, to the sample members already simulated. Simulated samples of any size can then be generated by running through the simulation any desired number of times. Each pass through the routine uses a new set of pseudorandom numbers to guide the generation of data according to the model which the investigator has laid down. (The term "pseudorandom" denotes numbers generated by an algorithm which produces a sequence of numbers which, while determinate, nevertheless possesses many of the most important characteristics of a sequence of random numbers.)

The simulation starts with the assumption that there is a single predictor variable which affects both an outcome variable (here, for simplicity in presentation, postulated to be income) and the likelihood of enlisting in the Service. While the analysis of data obtained from samples of populations of interest to the Army would require the use of a multivariate predictor variable which might include income, employment status, scholastic record, family military history, and other predictors, the present simulation will include only the single predictor. In any case, multiple predictors would usually be combined into a single linear combination which would act as the univariate predictor used here. For the purposes of this document, that predictor is called PRED and will be specified as being normally distributed with a mean of 0 and a standard deviation of 1, or $N(0,1)$. The use of the normal distribution is taken as an approximation to the conditions which are customarily approximated by data gathered from samples of populations of interest to the Army. Such simplifying assumptions as normality (multivariate or univariate) are customarily made in simulations because they match the assumptions on which the derivations of standard statistical techniques are based. Data based on the assumption of normality can be used with a wide variety of statistical techniques, including logistic regression. If, at a later time, it becomes desirable to investigate the robustness of the techniques (i.e. their sensitivity to violations of the assumptions of the analytical techniques), such investigations may be carried out by systematically violating the assumptions of the tests at the time the data are generated, as by introducing perturbations into the data which will yield nonnormally distributed data.

The method chosen for generating the normally distributed random deviates for the simulations is known as the method of direct generation. It is presented in Appendix A.

The simulation provides for a single dependent variable, called OUTCOME. Again, in the analysis of operationally derived data there would be more than one outcome variable, probably several intercorrelated variables such as employment status, income, education, life satisfaction measures, and health-related variables, but in this simulation OUTCOME is modeled to represent an approximation of income. OUTCOME depends on PRED thanks to two mechanisms. OUTCOME is the sum of two terms; one term is linearly related to PRED with a correlation which may be specified in the simulation, and the other term derives its influence from Military Service, which is itself related to PRED through a propensity score correlated with PRED to an extent which may also be specified in the simulation. The steps necessary for the generation of those two terms of OUTCOME are described next.

The first term of OUTCOME is derived directly from PRED by a linear function which incorporates the simulated correlation between the predictor and this term of the outcome. Because this term of the outcome does not include the effect of Service, it is thought of as the raw outcome, and called ROUT. ROUT is computed by first generating a normal deviate $N(0,1)$ which is correlated with PRED to the desired extent, and then modifying that deviate so that it will have the mean and standard deviation chosen by the investigator (the formal properties and statistical significance of any findings based on the outcome are not affected by the choice of the mean and standard deviation, but it is easier to understand the output of simulations when their units are plausible). For this simulation, define RPREDROUT as being the correlation between PRED and ROUT, ROUTMEAN as being the mean of ROUT, and ROUTSD as being the standard deviation of ROUT. If no member of the simulated sample served in the military, then this linear relationship would fully describe the relation between the predictor and the outcome. However, the presence of the second term in the simulation introduces the possibility of an influence of Military Service.

The second term of OUTCOME is the term due to Military Service. In the simulation, the magnitude of the effect of Service is modeled by a random normal variable called SER with a mean of SERMEAN and a standard deviation of SERSD. The magnitude of Service is not associated (correlated) with PRED, but whether or not one serves is influenced by PRED. Those simulees who are identified as serving have SER added to ROUT to compute OUTCOME, while those who do not serve do not have SER added to ROUT; in that case OUTCOME is simply equal to ROUT.

PRED acts to influence whether one serves in the military via a variable called the propensity score, here modeled by PROP, which is normally distributed with a mean and standard deviation specified at the time the simulation is run. As is the case with ROUT, PROP is linearly related to PRED. The correlation between the predictor (PRED) and the propensity score (PROP), called in the simulation RPREDPROP, is specified when the simulation is run. The mean and standard deviation of PROP are also specified for the simulation; they are governed by the variables PROPMEAN and PROPSD.

The propensity score, PROP, exerts its influence by governing the probability that a simulee serves. The probability of Service is given by the variable PSERVE, calculated by the relationship

$$PSERVE = 1/(1+\exp(-PROP)).$$

The intuitive or operational significances of most variables in the simulation are easily grasped based on a knowledge of basic regression and statistical analysis. The effects of PROPMEAN and PROPSD, however, are less intuitively obvious. The shape of the relationship with PROP on the x-axis and PSERVE on the y-axis is a logistic ogive which is asymptotic to a zero value of PSERVE when PROP is much less than 0, and asymptotic to 1 when PROP is much greater than 0. When the propensity score equals 0, then the probability of serving equals .5. Changing the mean of the propensity score distribution causes the average probability of serving to depart from .5, with increasing values of propensity being associated with increased probabilities of Service. The standard deviation of the distribution of propensity scores governs the steepness of the ogive at its steepest point, the point of inflection, where PSERVE equals .5. The larger values of PROPSD are associated with steeper curves, and so with distributions of PSERVE with relatively few values in the intermediate range and more values close to 0 and 1.

In the simulation, each simulee is designated as serving or not serving. The decision for each simulee is made by drawing a pseudorandom number from a uniform distribution on the interval 0,1. If the number is less than or equal to PSERVE, then the simulee is designated as serving, and the variable SERVE is given a value of 1. Otherwise SERVE is given a value of 0. In that way each simulee has a probability of serving equal to PSERVE, and the propensity score, PROP, had its intended effect.

Summary of Variables in the Simulations

Variables Calculated for Each Simulee

PRED	The predictor variable, normally distributed with a mean of 0 and SD of 1, i.e. $N(0,1)$
WV1 --WV3	Three deviates $N(0,1)$ which will be used in later steps of the simulation, called working variables (hence WV1, WV2, and WV3)
OUTCOME	The simulated outcome variable
ROUT	The "raw" outcome, or outcome without the effect of Service, correlated with PRED
PROP	The simulee's propensity to serve, correlated with PRED
PSERVE	The probability of serving, derived from PROP
SERVE	A binary variable specifying whether simulee served (1) or not (0)
SER	The magnitude of the effect of Service for a given simulee.

Variables Used in the Simulation and Specified by the User when Simulation is Conducted

ROUTMEAN	The mean of ROUT
ROUTSD	The standard deviation of ROUT
RPREDROUT	The correlation between PRED and ROUT
SERMEAN	The mean of the effect of Service
SERSD	The standard deviation of the effect of Service
RPREDPROP	The correlation between PRED and PROP
PROPMEAN	The mean of the propensity scores
PROPSD	The standard deviation of the propensity scores

Given that there are effects both of Service and of the predictor variables, the task of the propensity score analysis will be to determine the relative contributions of the two parts, the direct action of the predictors and the action through Service. The main difference between this simplified model and the operational multivariate model is that the multivariate model may (almost certainly will) have somewhat different sets of weights for predicting propensity and outcome from the predictor variables. This model might be improved by incorporating that feature, but it is at least partially modeled by allowing the correlation between predictors and raw outcome to differ from the correlation between predictors and propensity.

The analysis to be modeled (i.e. the one which would prevail in the analysis of operational data) requires that the propensity scores be estimated by logistic regression, then the cases sorted on the basis of propensity score, and then matching cases made. That analysis will be applied to the simulated data.

The simulation requires the following steps:

1. For each case to be simulated (i.e. each simulee), generate four uncorrelated normal random deviates, N(0,1). Designate these as the predictor, PRED, and three working variables WV1, WV2, and WV3.

2. Calculate the outcome variable without the effect of Service, the raw output variable. This is called ROUT.

$$ROUT = ((RPREDROUT * PRED + ((1-RPREDROUT^2)^{.5}))* WV1) * ROUTSD + ROUTMEAN$$

3. Use PRED and WV1 to generate a propensity score (PROP) correlated with the predictor according the specified RPREDPROP. Use PROPMEAN and PROPSD to adjust the size of the deviate so that its mean and SD will be appropriate. Use the formula

$$PROP = ((RPREDPROP * PRED + ((1-RPREDPROP^2)^{.5}))*WV2) * PROPSD + PROPMEAN.$$

4. Find each simulee's probability of serving (PSERV) by using the formula

$$PSERV = 1/(1+exp(-PROP))$$

5. Determine whether each simulee enlisted by selecting a random number on the uniform interval 0,1 and determining whether it is less than the probability of enlisting. If so, the simulee enlisted and the variable SERVE takes the value 1, otherwise the variable SERVE takes the value 0.

6. Determine the effect of Service by selecting a normally distributed variable from a population with specified mean and standard deviation and multiplying it by 0 or 1, depending on whether the simulee enlisted. That is called SER, and is uncorrelated with the other variables.

$$SER = WV3 * SERSD + SERMEAN.$$

7. Determine the final outcome by adding the two outcome variables, that due to the predictor alone and that acting through the enlistment variable. The Service-connected effect, SER, must be multiplied by SERVE to incorporate whether or not the individual served.

$$OUTCOME = ROUT + SER * SERVE$$

A sample simulation run for 100 simulees is given in Appendix B.

Data Sources

As indicated in the proposal for this project, one aim was to explore the availability of data for possible analysis in a possible Phase II. The project investigator met, as planned, with a project consultant who was believed to be fully familiar with potential sources of data which might be used to test hypotheses relevant to research in the area of effects of Service. This meeting was the occasion of a visit to the Washington, D. C. area.

THIS PAGE INTENTIONALLY LEFT BLANK

III. RESULTS

Simulations

The raw data of a representative simulation are provided in Appendix B. This set of data was generated in accordance with parameters which were estimated to be representative of those likely to be obtained from data gathered from populations of interest to the Army. A set of data prepared from the same input parameters was analyzed in three stages in order to determine whether the propensity score matching would yield an analysis of sufficient sensitivity to justify further investigation.

The first stage of the analysis was to account for the tendency of the simulees to enlist or not to enlist by means of logistic regression. A widely available logistic regression package was used for that analysis (Dallal, 1988)[11]. The analysis provided regression coefficients to be applied to the sample simulation, to give the prediction equation

$$CALCPROP = -.1848 \ PRED - .5013.$$

The equation indicates that a propensity score may be calculated (CALCPROP) for each member of the simulated sample by multiplying the PRED score for that simulee (i.e. the predictor variable) by -.1848 and adding the intercept of -.5013. A propensity score variable was then computed for each simulee and appended to the data available for each simulee.

The second stage of the analysis consisted in sorting all 100 simulees in decreasing order of calculated propensity scores, and then developing two matched samples. The matches samples were achieved by taking successive pairs of simulees, starting from the top of the order list, and determining whether each pair met both of two criteria. The first criterion required that one member of the pair had "served" and the other not had not "served" (i.e. one of the simulees having had a value of 1 for SERV, and the other having had a value of 0). The second criterion required that the two propensity scores for the members of the pair differ by no more than an absolute value of 0.02. Thus a sample was established in which propensity scores were matched for pairs of the sample, but one member of each pair of the sample had served, and one member had not served.

---

[11]Dallal, G.E. (1988). Logistic: A logistic regression program for the IBM PC. The American Statistician, 42: 272.

The above-described matching process required that a number of cases not be analyzed because suitable matching cases could not be found. In this particular analysis, 29 pairs survived the matching process, indicating that 58 of the original sample of 100 "survived" to give useful data. The remaining 42 cases could not be analyzed because suitable matching partners could not be found, either because there was no suitable matching propensity score or because there was an imbalance in the number of those who served and who did not serve.

The final stage of the analysis consisted of determining whether there was a detectable Service effect in the paired data. The pairs of data were analyzed by means of a correlated data t-test. The results of that analysis showed that the mean difference between those who served and those who did not serve was 571.65, and the standard deviation was 2657. The standard error of the mean was thus 493, resulting in a t-value of 1.16, which is not significant. As mentioned above, these units were arbitrary, and may be rescaled by linear transformation to units convenient to the user. In this case, the units were chosen to represent plausible dollar amounts. Thus in this simulation it was possible to detect a statistically insignificant difference of $571.65.

The difference between those who served and those who did not serve reflects the simulated value of 1600 (the value of SERMEAN, the mean effect of service, specified in Appendix B), and the standard deviation is also within expected limits. The insignificant t-test result may be due to nothing more than the sample size of 100 cases, not all of which could be analyzed because of lack of suitable matching candidates, but the parameters of the simulation were chosen to yield, in this initial analysis, positive results with small samples.

Data Sources

A half-day meeting with the project consultant revealed that his knowledge and expertise were more in the logistics of personnel research and the operations necessary to carry them out than in the details of the availability of data. In any case, the project Contracting Officer's Representative made available data which were available for use of the contractor. Repeated attempts to access these data with the facilities of several local institutions were not successful owing to unavailability of suitable software.

IV. DISCUSSION AND CONCLUSIONS

As seen in the results section, the calculations derived from the simulations do not appear to support the feasibility of using the propensity score methodology under the sample size investigated. There are three possible reasons for this. First, the methods may be insufficiently sensitive to detect differences of the magnitudes associated with reasonable estimates of effect size at the small sample size used. Extrapolation to a sample size of 1000 indicates that significant results would be obtained (increasing the sample by a factor of 10 would reduce the standard error of the mean by a factor equal to the square root of 10, thus the above-reported t value of 1.16 would be expected to increase to a significant 3.67, $p < .01$). This increase in the estimated significance is probably a conservative estimate because having a sample of 1,000 simulees, rather than 100, would allow more of the simulees to be successfully matched with others who met the matching criteria. In the case of the simulation reported here, 58% of the original sample "survived" to give useful data; with a sample ten times as large it is estimated that over 90% would survive. The increase in sample size then has two benefits -- one direct and one indirect. Other techniques to increase the yield of the simulations are possible. For example, the number of simulees can be increased to the point that the desired number of usable rather than total, cases is obtained. That procedure is justifiable in part because it is also used in gathering data for surveys in the conduct of field research, as when a population subgroup is oversampled to ensure a sufficient number of cases. Second, the method of deriving the simulations may be suboptimal, in spite of considerable attention and a number of alternative approaches having been considered. Third, the analysis of the simulated data may have contained unwarranted assumptions or less than optimal techniques. All of these possibilities remain under consideration. The devising, conduct, and analysis of the simulations proved less tractable than anticipated, however, which limited the number of simulations which could be conducted within the constraints of project resources.

The results of the simulations used for the present investigations remain somewhat enigmatic. It was anticipated that they would yield positive findings with modest sample sizes and that such findings would be usable to guide further investigation using both simulations and data from surveys of Service samples. While the findings do not yet appear to justify further investigation with operational data or, less still, with data gathered specifically to be used within the framework of propensity score theory, the question of the suitability of the method must remain open. The present investigator has reviewed the simulation and analyses in detail and believes them to be sound and that further investigation has a high likelihood of evolving a method which will be of use to the Army.

The current research was not able to reach conclusions which support the intended level of confidence in the method, but the method appear to be sufficiently promising to pursue in further research. It should be noted that since large-sample sizes characterize much of Military research in the behavioral and social sciences, the prediction of positive findings based on large samples may be relevant.

In short, the results of this analysis are not fully conclusive, but within the scope of the investigation it was not possible to determine conclusively the potential of the propensity score methodology. The question of the benefits of propensity score remains open, but cautious optimism is appropriate.

REFERENCES

Abramowitz, M. & Stegun, I. A., Eds. (1972). <u>Handbook of mathematical functions.</u> Originally published by National Bureau of Standards. Reprinted: New York: Dover Publications.

Brumagim, A.L. & Daymont, T. N. (1989). The role of military educational benefit programs: Impacts on minority opportunities. <u>Industrial Relations Research Association 42nd Annual Proceedings.</u> 315-325.

Crane, J. R. & Wise, D. A. (1987). Military Service and the civilian earnings of youths. In D. A. Wise (Ed.), <u>Public Sector Payrolls.</u> Chicago: University of Chicago Press. 119-137. (as cited in Lakhani, 1994).

Dallal, G. E. (1988). Logistic: A logistic regression program for the IBM PC. <u>The American Statistician, 42</u>, 272.

Daymont, T. N. & Andrisani, P. J. (1986). <u>The economic returns to Military Service</u>. (Technical report USARECSR 86-11) Fort Sheridan, Illinois: U.S. Army Recruiting Command.

Lakhani, H. (1994). <u>The socioeconomic effects of Military Service: Reserve/Guard</u>. Paper presented at the 69th annual Western Economic Association International Conference, Vancouver, Canada, 1994. (Preliminary version).

Laurence, J. H., Ramsberger, P. F., & Gribben, M. A. (1989). <u>Effects of military experience on the post-Service lives of low-aptitude recruits: Project 100,000 and the ASVAB misnorming</u>. (Technical Report 89-29) Alexandria, VA: Human Resources Research Organization.

Martindale, M. & Poston, D. L. (1979). Variations in veteran/nonveteran earnings patterns among World War II, Korea, and Vietnam War cohorts. <u>Armed Forces and Society, 5</u>, 219-243.

Rosenbaum, P. R. & Rubin, D. R. (1984). Reducing bias in observational studies using subclassification on the propensity score. <u>Journal of the American Statistical Association, 79</u>, 516-524.

Sticht, T. G. (1992). <u>How Military Service helped low-aptitude, economically disadvantaged young men of the mid-1960's escape poverty</u>. Research note Number 1. San Diego: Applied Behavioral and Cognitive Sciences, Inc.

Sticht, T. G., Armstrong, W. B. Hickey, D. T., & Caylor, J. S. (1987). <u>Cast-off youth: Policy and training methods from the military experience.</u> New York: Prager.

Villamez, W. J. & Kasarda, J. D. (1976). Veteran status and socioeconomic attainment. <u>Armed Forces and Society, 2</u>, 407-420.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A

The Generation of Normally Distributed Pseudorandom Numbers

The generation of normally distributed pseudorandom numbers is a well-studied problem with several satisfactory solutions (Abramowitz & Stegun, p. 952).[12]. For the simulations undertaken for this inquiry, the method known as direct generation was used. Direct generation requires first the generation of two pseudorandom numbers from a uniform distribution on the interval 0,1. The two pseudorandom numbers, here designated as U_1 and U_2, are then transformed to normal deviates with mean 0 and standard deviation 1, D_1 and D_2, by means of the formulas

$$D_1 = (- 2 \ln U_1)^{1/2} \cos 2 \, pi \, U_2$$

$$D_2 = (- 2 \ln U_1)^{1/2} \sin 2 \, pi \, U_2.$$

Such deviates may be transformed to deviates with desired mean and standard deviation by multiplying by the standard deviation and adding the mean. When correlated deviates are required, as for the simulations in the present inquiry, then D_2 is modified (Abramowitz & Stegun, 1972) according to the following formula, where r = desired correlation coefficient.

$$D_2' = r \, D_1 + (1 - r^2)^{1/2} D_2 \, .$$

---

[12]Abramowitz, M. & Stegun, I.A., Eds. (1972). Handbook of mathematical functions. Originally published by National Bureau of Standards. Reprinted: New York: Dover Publications.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B

Sample Simulation

This appendix presents an example of a simulation based on the method outlined in the text of this report. Values given for the simulated variables were chosen to approximate those expected to be found in populations of interest.

The following values are selected to be used in this sample run of the simulation:

ROUTMEAN	The mean of ROUT	16,000
ROUTSD	The standard deviation of ROUT	2,000
RPREDROUT	The correlation between PRED and ROUT	0.3
SERMEAN	The mean of the effect of Service	1,600
SERSD	The std. dev. of the effect of Service	200
RPREDPROP	The correlation between PRED and PROP	0.4
PROPMEAN	The mean of the propensity scores	0
PROPSD	The std. dev. of the propensity scores	1

The following variables are calculated during the simulation:

PRED	the predictor variable
WV1,2,3	intermediate working variables
OUTCOME	final outcome variable
PSERVE	probability of serving
SERVE	Service: 1=served, 0=not served
SER	magnitude of effect of Service

PRED	WV1	WV2	WV3	PROP	p(SERV)	SERV	ROUT	SER	OUTCOME
-1.36	-1.998	1.82	-0.76	-2.314	0.090	0	18251	1448	18251
-0.53	-2.092	0.12	0.69	-2.153	0.104	0	15791	1738	15791
-1.42	-1.749	-0.21	-0.52	-2.094	0.110	0	14481	1496	14481
-0.43	-2.000	1.55	-1.31	-2.037	0.115	0	18490	1338	18490
-0.73	-1.584	0.54	-0.23	-1.731	0.151	0	16401	1554	16401
-1.00	-1.469	-1.04	1.50	-1.700	0.154	0	13292	1899	13292
0.49	-1.880	0.36	-1.28	-1.645	0.162	0	17052	1344	17052
-0.63	-1.362	1.32	0.01	-1.488	0.184	1	17909	1601	19510
0.55	-1.723	0.98	0.53	-1.478	0.186	0	18238	1706	18238
0.02	-1.503	0.80	0.86	-1.427	0.194	0	17485	1773	17485
-0.17	-1.372	1.96	-0.22	-1.359	0.204	0	19466	1557	19466
-1.39	-0.971	1.71	-0.88	-1.343	0.207	0	18033	1425	18033
-0.87	-1.111	1.15	0.28	-1.321	0.211	0	17412	1655	17412

-1.21	-1.000	-0.40	1.30	-1.317	0.211	0	14299 1859	14299
0.02	-1.287	-0.32	0.61	-1.222	0.228	0	15422 1722	15422
-1.65	-0.748	-0.14	-0.66	-1.210	0.230	0	14419 1468	14419
-0.12	-1.156	0.19	-0.47	-1.138	0.243	0	16262 1506	16262
-0.29	-1.078	0.39	-0.90	-1.116	0.247	0	16484 1420	16484
-0.68	-0.908	-0.01	-0.24	-1.071	0.255	0	15430 1551	15430
-0.94	-0.800	1.31	-0.24	-1.044	0.260	0	17644 1551	17644
-1.08	-0.640	-0.08	1.67	-0.935	0.282	1	14987 1935	16922
0.42	-1.107	1.18	-1.60	-0.931	0.283	1	18503 1280	19783
-1.33	-0.458	-1.42	-0.20	-0.836	0.302	0	12325 1561	12325
-0.10	-0.817	0.10	1.78	-0.810	0.308	1	16103 1956	18058
-1.46	-0.369	1.77	2.49	-0.788	0.312	1	18071 2098	20170
-0.54	-0.626	0.24	-0.44	-0.758	0.319	1	16011 1512	17523
-0.05	-0.742	1.19	1.25	-0.721	0.327	0	18141 1849	18141
0.40	-0.876	-0.15	-1.49	-0.715	0.328	1	16052 1301	17353
1.93	-1.316	-1.04	-0.02	-0.675	0.337	0	15640 1596	15640
1.22	-0.968	-1.11	1.32	-0.559	0.364	0	14944 1863	14944
-1.16	-0.216	-1.42	-0.56	-0.555	0.365	1	12460 1487	13948
1.00	-0.877	-0.33	0.92	-0.537	0.369	1	16195 1785	17980
-0.57	-0.349	0.81	-1.46	-0.504	0.377	0	17033 1308	17033
-0.09	-0.448	-1.90	-0.64	-0.455	0.388	0	12441 1472	12441
0.08	-0.465	0.34	1.12	-0.421	0.396	1	16690 1824	18515
-0.76	-0.184	-0.38	0.38	-0.403	0.401	0	14700 1676	14700
-1.41	0.081	-0.88	0.09	-0.345	0.415	0	13253 1617	13253
-1.91	0.259	1.15	-0.23	-0.327	0.419	0	16581 1554	16581
-1.12	0.012	0.27	0.30	-0.323	0.420	1	15596 1660	17256
-0.55	-0.145	0.84	-0.50	-0.304	0.424	0	17103 1501	17103
-0.33	-0.131	0.17	0.42	-0.225	0.444	0	16039 1684	16039
-1.49	0.283	-0.12	-0.06	-0.177	0.456	1	14586 1589	16175
-0.52	0.005	0.43	1.07	-0.150	0.462	0	16371 1814	16371
-0.77	0.088	-0.13	1.19	-0.149	0.463	1	15135 1839	16974
-0.35	-0.046	-0.46	-0.03	-0.148	0.463	1	14887 1594	16481
1.77	-0.709	0.61	-0.55	-0.146	0.464	0	18535 1490	18535
0.21	-0.211	-0.29	-0.39	-0.137	0.466	1	15639 1523	17162
-0.28	-0.007	-0.40	0.22	-0.091	0.477	1	15048 1645	16693
-0.45	0.063	-0.29	0.34	-0.075	0.481	0	15109 1667	15109
-0.26	0.056	-0.09	-0.28	-0.024	0.494	0	15631 1544	15631
1.92	-0.606	0.24	-0.62	-0.001	0.500	1	17972 1476	19448
0.48	-0.145	-1.23	0.23	0.006	0.502	1	14127 1647	15774
-0.34	0.134	0.78	0.09	0.027	0.507	1	17164 1619	18783
2.15	-0.614	0.01	0.70	0.061	0.515	1	17740 1739	19480
-0.43	0.202	-1.94	1.25	0.063	0.516	0	12103 1851	12103
0.51	-0.048	-0.72	-0.44	0.108	0.527	1	15080 1513	16593
0.29	0.024	0.98	1.35	0.110	0.528	0	18027 1869	18027
-0.43	0.261	0.27	-0.93	0.119	0.530	0	16149 1413	16149
0.57	-0.052	-0.04	0.86	0.120	0.530	0	16378 1773	16378
0.78	-0.112	-0.22	1.39	0.127	0.532	1	16230 1878	18108
1.11	-0.201	0.77	0.72	0.140	0.535	0	18305 1744	18305
1.52	-0.331	0.75	0.08	0.141	0.535	1	18586 1615	20202
0.03	0.182	0.82	-0.31	0.183	0.546	1	17529 1538	19067
-1.21	0.575	-0.83	-1.41	0.186	0.546	1	13505 1317	14822
-0.25	0.301	0.83	0.08	0.211	0.553	1	17320 1616	18936
0.42	0.099	-0.06	0.08	0.219	0.555	1	16231 1616	17847
-0.58	0.423	-0.48	1.23	0.229	0.557	1	14654 1846	16500
0.13	0.238	0.24	-0.31	0.265	0.566	1	16550 1537	18087

0.03	0.270	-0.68	0.83	0.267	0.566	0	14779	1767	14779
-0.54	0.467	-0.76	1.69	0.283	0.570	1	14179	1937	16116
0.69	0.079	-1.55	0.55	0.284	0.570	0	13713	1709	13713
1.59	-0.139	2.64	1.43	0.344	0.585	0	22106	1885	22106
-0.18	0.452	0.35	-0.29	0.378	0.593	0	16509	1543	16509
-1.09	0.752	1.21	0.05	0.389	0.596	1	17346	1611	18956
0.64	0.289	1.03	1.04	0.466	0.615	0	18395	1809	18395
0.63	0.308	0.44	-0.35	0.483	0.618	1	17309	1529	18838
0.09	0.584	-0.67	2.04	0.585	0.642	1	14841	2008	16849
-0.11	0.680	-0.41	-0.48	0.615	0.649	1	15149	1504	16653
1.37	0.347	-0.90	0.25	0.741	0.677	0	15449	1651	15449
-0.21	0.896	0.01	-0.16	0.792	0.688	0	15856	1568	15856
0.63	0.642	0.38	1.52	0.802	0.690	0	17198	1904	17198
1.27	0.454	1.58	1.12	0.813	0.693	1	19912	1825	21736
-0.34	0.991	0.68	0.73	0.843	0.699	1	16968	1746	18714
0.44	0.772	0.85	-0.15	0.869	0.705	1	17921	1569	19490
-0.17	0.995	-0.70	-0.61	0.899	0.711	0	14589	1477	14589
0.22	0.967	0.21	0.50	0.988	0.729	1	16561	1701	18262
0.89	0.825	0.97	-1.90	1.053	0.741	1	18483	1221	19703
1.35	0.705	-0.60	1.05	1.076	0.746	1	15969	1809	17779
1.13	0.782	1.43	0.45	1.084	0.747	1	19523	1689	21212
1.14	0.808	-2.56	-0.29	1.113	0.753	0	12224	1543	12224
1.11	0.820	1.68	-0.50	1.114	0.753	0	19957	1499	19957
0.00	1.169	-0.13	1.17	1.117	0.753	1	15757	1835	17592
-0.63	1.460	-1.22	0.68	1.202	0.769	1	13253	1736	14989
1.86	0.717	-0.61	0.15	1.241	0.776	1	16378	1631	18009
1.55	0.817	0.04	0.63	1.244	0.776	1	17314	1726	19040
-1.23	1.753	1.34	0.78	1.304	0.786	0	17472	1756	17472
1.57	0.873	0.44	-0.06	1.305	0.787	1	18066	1587	19653
-0.32	1.683	0.45	-0.81	1.511	0.819	0	16569	1439	16569
0.10	1.747	0.42	0.20	1.698	0.845	1	16850	1641	18491
0.63	1.592	1.35	0.56	1.707	0.846	0	18981	1711	18981

Means and Standard deviations of the values associated with the above 100 simulees:

	PRED	WV1	WV2	WV3	PROP	p(SERV)	SERV	ROUT	SER	OUTCOME
Mean	-0.011	-0.128	0.18	0.20	-0.125	0.476	0.47	16313	1640	17092
S.D.	0.928	0.885	0.94	0.87	0.928	0.199	0.50	1859	174	2046